Movement Disorders (Wiley Online Library): Review (full paper)

DOI: 10.1002/mds.26718

Accepted 13 June 2016

Free-living monitoring of Parkinson's disease: lessons from the field

Silvia Del Din, PhD[1], Alan Godfrey, PhD[1], Claudia Mazzà, PhD[2,3], Sue Lord, PhD[1], Lynn Rochester, PhD[1]

[1] Institute of Neuroscience | Newcastle University Institute for Ageing, Clinical Ageing Research Unit, Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne, NE4 5PL, UK

[2] Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK

[3] INSIGNEO Institute for *in silico* medicine, The University of Sheffield, Sheffield, UK

Corresponding author:

Lynn Rochester PhD

Clinical Ageing Research Unit,

Campus for Ageing and Vitality,

Newcastle University,

NE4 5PL,

UK

Email: lynn.rochester@ncl.ac.uk

Phone: +44 (0) 191 208 1291

Word Count (excluding abstract, legends and references): 4657

Abstract word count: 249

Running title (not exceeding 45 letters and spaces): Wearable technology for Parkinson's disease

Key words (up to 5): Wearable technology, Parkinson's disease, remote monitoring, free-living assessment

Figures: 3

Tables: 2

Conflict of Interest: The authors declare that they have no conflict of interest.

Abstract

Wearable technology comprises miniaturized sensors (e.g. accelerometers) worn on the body and/or paired with mobile devices (e.g. smart phones) allowing continuous patient monitoring in unsupervised, habitual environments (termed free-living). Wearable technologies are revolutionising approaches to healthcare due to their utility, accessibility and affordability. They are positioned to transform Parkinson's disease (PD) management through provision of individualised, comprehensive, and representative data. This is particularly relevant in PD where symptoms are often triggered by task and free-living environmental challenges that cannot be replicated with sufficient veracity elsewhere. This review concerns use of wearable technology in free-living environments for people with PD. It outlines the potential advantages of wearable technologies and evidence for these to accurately detect and measure clinically relevant features including motor symptoms, falls risk, freezing of gait, gait, functional mobility and physical activity. Technological limitations and challenges are highlighted and advances concerning broader aspects are discussed. Recommendations to overcome key challenges are made. To date there is no fully validated system to monitor clinical features or activities in free living environments. Robust accuracy and validity metrics for some features have been reported, and wearable technology may be used in these cases with a degree of confidence. Utility and acceptability appears reasonable, although testing has largely been informal. Key recommendations include adopting a multi-disciplinary approach for standardising definitions, protocols and outcomes. Robust validation of developed algorithms and sensor-based metrics is required along with testing of utility. These advances are required before widespread clinical adoption of wearable technology can be realised.

Introduction

Wearable technology and connected devices (WTCD) are positioned to become ubiquitous in research and healthcare settings. WTCD comprise electronic technology worn on the body or embedded into mobile phones, watches, bracelets, and clothing, amongst others. The generic appeal of WTCD is obvious. Patient monitoring is free from contextual or environment barriers making assessment at home and in the community over continuous time periods (termed free-living) feasible and ecologically valid [1]. Moreover data are free from the confounds of observer bias and attentional compensation associated with a one off testing session under observation [2], while devices are relatively low cost making their use economically as well as practically feasible.

The benefits of remote monitoring with WTCD are multi-fold. Clinically, continuous monitoring of symptom severity and therapeutic response provides nuanced assessment. A complete picture of disease burden is available both to the clinician and the patient incorporating a broad range of features from the '*micro*' level of detail (e.g. disease symptoms, medication response and fluctuations, gait characteristics, turning, frequency of falls) through to more '*macro*' levels (e.g. habitual patterns of walking/activity, inactivity and sleep) (Figure 1). Enriched measurement, coupled with ease of use, also has implications for industry, paving the way for identification of early disease with the potential for enhanced diagnostic and progression markers (fundamental for trials of novel therapeutics and disease modifying therapies), harmonisation of outcomes and standardized testing protocols to enhance recruitment and assessment of treatments in clinical trials. For the patient, WTCD offer insight into symptoms, therapeutic efficacy and habitual mobility in the context of everyday life contributing to enhanced self-management that is both bespoke and contextualised.

Despite the recent explosion of low cost commercially available devices (for the general population) promoting personal monitoring and feedback, the application of WTCD in healthcare has not yet been established [3]. The lure of utility (i.e. ease of use, broad application, and low cost) is strong; however standards for clinical adoption and research application are far higher. While technology and design have advanced, algorithm development and data analysis have not kept pace. Validity and reliability are paramount and inform accurate detection and monitoring of disease and this next step is critical before widespread adoption [4]. Although there are promising signs, there is still

no single system/gold standard being used for remote monitoring [5, 6]. Therein lies both the opportunity and the challenge.

This paper considers issues related to free-living monitoring from predominantly single sensor-based devices (e.g. accelerometers and gyroscopes). We examine the ability of WCTD algorithms to accurately detect a range of clinical features and report on criterion and discriminative validity of outcomes derived from WCTD. Utility and feasibility are also considered. Clinical features include monitoring of motor symptoms, medication response, sleep, falls and falls risk, freezing of gait (FOG), gait, functional mobility and physical activity (ambulatory activity and sedentary behaviour). This rapidly expanding field and has been the subject of a number of recent systematic reviews [7-9] including Sánchez-Ferro et al. within this issue to which the reader is referred. We have therefore adopted a broader approach and provide a structured overview of the current status of continuous patient monitoring in the home and community in Parkinson's disease (PD) which we define as 'free-living'. We address four key aims: (1) the role and benefits of free-living monitoring; (2) the validity and utility (acceptability and feasibility) of WTCD to monitor a range of key clinical features relevant to PD; (3) critical challenges for adoption of WTCD for free-living assessment; and (4) future developments in this rapidly developing field. Throughout we focus mainly on the application of passive (no interaction from patient) single sensor-based devices and their application in PD but where relevant draw from work in ageing cohorts. Finally, we make recommendations based on this overview to progress free-living monitoring in PD.

Does free-living monitoring confer an advantage over clinical assessment in PD?

Due to its heterogeneity and complexity, clinical assessment of PD is challenging. The intrinsic, fluctuating nature of PD and biphasic medication response in advanced disease requires continuous evaluation over prolonged periods to gain an accurate picture of symptoms and their fluctuations. The influence of attention on performance is well recognised especially with symptoms such as FOG, leading to an inaccurate clinical picture [2, 8]. Assessments requiring concentration and recall such as falls diaries are further compromised by cognitive impairment, thus limiting utility. Also, use of clinical scales is restrictive. The Unified Parkinson's Disease Rating Scale, (UPDRS) [10],

110 although highly relevant to PD, is dependent on the patient's status at the time of assessment and
111 limited by subjectivity and clinical expertise. WTCD overcome many of these limitations by
112 objectively quantifying clinically relevant outcomes. Variation in testing is reduced [3, 11, 12]. Patients
113 also have much to gain from this approach, with less emphasis during clinical visits on symptom
114 recall and evaluation of therapeutic response. Continuous monitoring also provides greater potential
115 for patient involvement in defining optimal management [12].

116 Measurement with WTCD is diverse. A single WTCD has the potential to provide the
117 clinician/researcher with a comprehensive picture of their patient within one assessment. For example,
118 Figure 1 shows that placement of a single sensor can quantify features such as volume and pattern of
119 habitual behaviours (e.g. walking, sleeping, sedentary time, Figure 1, A) (defined here as *macro*). The
120 raw signal (Figure 1, B) can then be further broken down to detect very discrete features (e.g. a fall,
121 gait characteristics, turning and freezing, figure 1, C-H) (defined here as *micro*). Taking this approach
122 enables multi-level measurement [13].

123 <Figure 1>

124

125 **Free-living assessment of clinically relevant features in PD: a valid alternative to conventional**
126 **clinical assessment?**

127 Despite the obvious advantages of free-living assessment an important question remains – are
128 the outcome measures derived from WTCD suitable for current clinical use and will patients and
129 professionals use WTCD? Table 1, which form the basis of this section, provides an overview of
130 detection accuracy, validity and utility of some WTCD. Our main inclusion criterion was that WTCD
131 had been applied to free-living monitoring under either totally unsupervised or scripted protocol
132 conditions, with an exception made for studies where tests are conducted in formal settings to
133 optimise validation, such as detection of FOG. We report *criterion validity* from studies that examine
134 the association between WTCD-derived outcomes and other measures such as clinical scales. We also
135 report studies that test *discriminative validity*, which we define as the ability of WTCD-derived
136 outcomes to discern groups or phenotypes. The list is by no means exhaustive but provides a current
137 overview and highlights the vast interest in the area. We do not review static postural control despite

its obvious relevance to PD [14, 15], because studies are laboratory and/or clinic based, however, facets of postural control (e.g. dynamic, turning) are considered.

Motor symptoms, medication response and sleep. Continuous monitoring has a lot to offer over snapshot clinical assessments which may not reveal the true extent of symptom burden. Earlier use of WTCD for motor symptom measurement focused on evaluation of a single symptom to detect hypokinesia, dyskinesia, tremor, bradykinesia, and akinesia derived on/off medication status [16, 17]. This has evolved to assessment of multiple motor symptoms using either a single [18-20] or multiple sensor systems [17, 21-24]. To date preliminary results are promising. Overall, motor symptom measurement using WTCD is accurate and comparable with more established methods with some aspects of validity tested. Criterion validity is established for most motor symptoms (tremor, bradykinesia, dyskinesia) showing moderate to high correlations overall ($R > 0.65$) with standard clinical scales (e.g. UPDRS, Abnormal Involuntary Movement Score (AIMS), Modified Bradykinesia Rating Scale (MBRS), etc.) (see Table 1 for references). Measures of bradykinesia also show high specificity (88%) and sensitivity (95%) when compared to standardised tests (e.g. the Dot Slide test) [18]. Studies that test discriminative validity are not as advanced, apart from the work by Horne et al. which discerns motor symptom fluctuations in early stages of PD [20]. Single sensors are sufficiently robust for application, although there are question marks over aspects of utility for some systems which require technical mastery and are demanding on the user (see 'Utility' section). Whilst there have been a number of key developments in this area with motor symptom monitoring assessed at home, the test protocols are still largely controlled and scripted as highlighted in table 1. True passive monitoring without patient input is as yet an area to be developed but remains the area of greatest interest as it will give the most ecologically valid picture of motor symptom burden and therapeutic efficacy. Assessment of sleep also shows promise. WTCD-derived outcomes for sleep discriminate PD from older adults (OA) [25, 26] for *macro* outcomes (e.g. number and size of movements) with people with PD also showing increased episodes of nocturia, fewer turns during sleep, and greater arm movements.

Falls and falls risk. Accurate detection of falls and falls risk (ideally before the first ever fall) would greatly inform clinical management and therapeutic development and WTCD has a role to play. Real-world detection of falls however is technically challenging. A plethora of algorithms, devices, and device locations (chest, waist or wrist [27-31]) have been tested to improve the accuracy of falls detection, however, studies are almost completely limited to controlled settings and conducted on young healthy adults. Kangas et al. provides a rare example of using WTCD for falls detection in the real-world where falls were measured in institutionalised OA and verified by an observer [32]. Fall detection sensitivity was 80% with a falls alarm rate per hour of 0.025, denoting one false alarm over 40 hours of recording. This points to high accuracy, although the testing environment was far removed from 'free-living', and generalisability is therefore weak. Application in PD remains an area of unmet need. An alternative approach is to predict falls risk using WTCD which, in contrast to falls detection, is a more advanced field for both older adults and PD. Moreover, addressing a falls prevention approach could be argued to have greater clinical relevance [33, 34]. Studies have compared groups with and without falls in PD using free-living monitoring over 3-7 days. Falls risk factors derived from gait during free-living walking bouts [33, 34] were superior to laboratory-based gait speed and fall history to discriminate fallers from non-fallers [35-38]. Discriminative validity has been established for both *macro* and *micro* characteristics of gait and sedentary behaviour (Figure 1, A-B) which are associated with type of PD fallers [39] and fall history (fallers vs. non-fallers) in OA [38, 40] and PD [41], respectively. *Micro* features may offer more than *macro* features [36, 37], and contribute substantially to predicting falls both in fallers and non-fallers [37, 38]. Further refinement of algorithm and system development is however required to take the field forward.

Freezing of gait. Gait disturbances such as FOG are notoriously difficult to replicate in a controlled environment because of its spontaneous nature and the non-specific and poorly understood triggers that provoke it [3]. Clinical scales such as the UPDRS and NFOG [42] are subjective and therefore limited. Despite the obvious need, free-living monitoring of FOG in PD has not been achieved. Detection of FOG episodes has been tested in controlled and structured conditions where FOG is

provoked during the 'off' condition, using either timed-up-and-go (TUG) [43] or walking tasks. [44] Studies show high sensitivity (range: 84.3%-86.2%) and moderate to high specificity (range 66.7%-98.74%) for detection of FOG, and moderate agreement with clinical measures [43, 44]. These results provide a critical step from which validation can be extended to free-living. An alternative approach is to identify potential predictors of FOG to understand the mechanisms and target therapeutic developments. A recent study comparing freezers vs. non-freezers found frequency-based gait characteristics collected during 3 days of free-living discriminated freezers. Gait characteristics were also moderately correlated with clinical measures of FOG [45]. Further work is needed before free-living monitoring can be used for FOG detection or indeed for understanding the characteristics of FOG but initial results are promising.

Gait. Measurement of gait per se (*micro* characteristics - Figure 1, E-F) is also of interest to the clinician to evaluate efficacy of clinical management (due to dopa-resistance) as well as for its potential for use of discrete gait characteristics as diagnostic, prognostic and progression markers [46-48]. Gait assessment during free-living assessment also captures ongoing environmental and cognitive challenges which impair gait performance. Assessment in this context has greater ecological validity and gives a true picture of the burden of disease [3, 7, 49]. Algorithms have been validated to detect discrete gait characteristics in the laboratory and also in proxy validation studies [50-55]. Results showed good agreement with trusted gold standard reference (e.g. GaitRite or optical motion capture systems) for the majority of gait characteristics with potential advantages for asymmetry and variability measures. Apart from Del Din et al. [49], the few studies that have examined gait in free living conditions, quantify few gait characteristics [56-61]. Discriminative validity has been tested, and has been shown to discriminate between PD and OA [49, 57], phenotypes of PD [61] and PD with higher or lower cognitive functions [60]. Aside from studies exploring falls and FOG risk highlighted previously [57] only one study has investigated the effect of environment on gait. Free-living gait characteristics showed better discriminative validity than those collected in the laboratory, especially for medium to long bouts [49]. Although initial work is promising, future work is required to confidently realise continuous monitoring of gait. There are also some fundamental challenges to the field (outlined below).

Measures of functional mobility. Tests of functional mobility such as turning and Timed up and Go (TUG) [62-64] measure combined movements that invariably incorporate postural transitions. Detection of movements during functional mobility tasks appears accurate [62, 63, 65], and validity (criterion and discriminative) has been established by a limited number of studies [62, 65]. Mean turn velocity, slower walking and turning, shorter steps and lower cadence distinguished PD from controls [62, 64] and also showed greater sensitivity to dysfunction than clinical rating scales [64, 65]. Of interest, free-living assessment appears to discriminate pathology better than testing in the laboratory [54] (Figure 1, G). Measurement of functional mobility tasks can therefore be undertaken with a degree of confidence during a standardised test at home, although further work is required to replicate these findings.

Ambulatory activity and sedentary behaviour. One of the earliest applications of WTCD aimed to quantify physical activity (e.g. ambulatory activity) amid rising concerns of the negative effects of sedentary behaviour on well-being. This is particularly relevant for people with PD because of the beneficial health benefits activity confers, and its role in mitigating secondary deficit. Ambulatory activity provides a picture of the true burden of disease and therapeutic efficacy [66]. Proxy measures such as activity logs and diaries are unreliable and lack responsiveness compared with continuous WTCD monitoring [67]. Physical activity such as intensity of movement (energy expenditure), temporal periods (bouts) of ambulatory activity (e.g. bouts of walking) and sedentary behaviours are quantified, from which *macro* outcomes can be derived [66, 68-70] (Figure 1, A-B). The field has advanced further with the application of non-linear approaches to data analysis which in some instances are more sensitive than measures of central tendency (Table 1, Figure 2), such as pattern (alpha (α)) rather than volume of sedentary behaviour showing discriminative properties [71]. Ambulatory activity differentiates disease stage [66], and progression [72, 73] and shows increased sensitivity to intervention [68, 74]. Rochester et al. [68] demonstrated the advantages of WTCD versus clinical measures when examining the impact of deep brain stimulation (DBS) on ambulatory activity. Whilst standard clinical measure for gait speed (4 meter test), levels of activity (Nottingham extended activities of daily living index (NEADL)) and disease progression (Hoehn and Yahr) failed to show the positive

effects of DBS on the outcomes, WTCD-based measures demonstrated significantly improved patterns of daily activity. Use of WTCD to measure ambulatory activity and sedentary behaviour is the most advanced of all the fields discussed in this section, and the most widely adopted. Nonetheless there are still questions over its application, driven by lack of common definitions of ambulatory activity, validation procedures and structured protocols in controlled settings for validation of algorithms [6]. These will be considered below.

Utility and feasibility of WTCD: how acceptable are they? Most studies do not intentionally test the feasibility and utility of WTCD but instead draw on secondary data such as informal comments from patients, reporting adverse events, data loss, or attrition in sensor use over the study period. Importantly, there are no overwhelmingly negative reports, suggesting that WTCD are broadly accepted. Although few studies have intentionally tested utility (which we describe as 'formal testing' in Table 1), some focused efforts have been made. Utility has been tested for wearable systems comprising interactive [75] or multiple sensors [17, 22, 23, 76], using both non-standardised and standardised questionnaires and rating scales[23] (e.g. the post-study usability questionnaire), comfort [75, 76] (e.g. comfort rating scale (CRS)) and 'wearability'/exertion [76] (e.g. Borg CR-10 Scale, Rapid Entire Body Assessment (REBA)). Overall the response has been positive, with WTCD generally well tolerated, comfortable and easy to use. Compliance is high, although in some cases results were influenced by socio-cultural aspects which may have positively biased results [23].

In summary, to date there is no fully validated WTCD system for continuous monitoring of patient clinical features. Overall, studies are small, there is no consistent reporting of outcome measures, protocols differ, and devices differ along with device placement. Comparison to a gold standard is difficult. Knowledge on patient acceptability is limited. A clear process for validation including replication in external data sets is essential with appropriate reporting according to a standard. However the WTDC community is aware that this is an important and emerging area of research with potential for high clinical uptake, and collaborative efforts are underway to redress these issues (see reviews [7-9]). Challenges to implementation are due at least in part to broader technological and

practical concerns which are common to all WTCD and influence their state of readiness, irrespective of application and use. Until these fundamental issues are redressed, robust use of WTCD will be compromised. The next section highlights some of these broad concerns and discusses approaches to advance the field.

Challenges to clinical adoption

We address 3 key areas fundamental to the use of WTCD that apply to all areas of measurement: (i) clear definitions of the clinical feature of interest, (ii) validation of real-world data and WTCD technical challenges, and (iii) consensus on outcomes. We illustrate these using examples from our own experience in gait and activity and that of others (Figure 3). Finally we summarise challenges with recommendations for future work and practical suggestions to inform the interested user (Table 2).

Defining the clinical feature. Although on the face of it this seems simple, there are many examples where unclear definitions have led to inconsistencies in outcomes and confusion when comparing between studies. A good example relates to ambulatory activity, from which *macro* (e.g. walking bouts) and *micro* level gait outcomes are derived that underpin many different clinical and research questions (Figure 1). This stems from a basic definition of what constitutes a walking bout. In some studies only purposeful bouts of walking are considered (with a cut-off threshold > 60 seconds) because regular steady state is more likely to be achieved, thus avoiding potential errors in misclassification from short bouts. However this is problematic because adults perform almost 90% of walking bouts in less than 60s [40, 49, 77] resulting in significant data loss and potentially missing the most relevant data (such as change in variability of walking pattern). Another approach is to include all bouts of walking [49] which is arguably more relevant in patient populations. However this is not a complete solution because disagreement also exists regarding the number of steps required for a bout, which may vary, ranging from >3 steps to >10 steps. As a consequence comparison across studies is impossible where difference in step counts range from 2,000 to 10,000 steps [66, 68, 72, 73]. The situation is further complicated by the use of ‘ghost’ (unknown to the end user and hard-wired into WTCD)

thresholds used by the manufacturer to define consecutive bouts of walking that have a major impact on *macro* outcomes [78] (e.g. total number and pattern of walking bouts) (Figure 3, (1)). This uneven approach significantly impacts on both *macro* and *micro* outcomes and therefore consensus as to a clear definition of walking is urgently required [6, 78]. Attempts are underway to improve definitions which will greatly help (Chastin et al.: AlPHABET: Development of A Physical Behaviour Taxonomy with an international open consensus[1]).

Algorithm development, validation and technical challenges: Influence of context and protocol. Establishing a gold standard to test algorithm validity for the range of features highlighted in this review during continuous uncontrolled monitoring in a free-living environment is a major challenge without obvious solutions. Real-life is unpredictable and unstructured. For example, context (environment and task) affects walking speed and direction which has implications for accuracy of algorithms used to detect steps and phases of the gait cycle from which gait characteristics are determined (Figure 3). Studies often adopt a number of different testing protocols and various sensor configurations (type and location (upper or lower body, Table 1) which also impacts the signal waveform influencing the accuracy of the algorithm used to extract micro outcomes and other type of information (features, outcomes). Moreover algorithms are usually validated using healthy controls data and adopted for analysing other groups' data (i.e. PD) without considering that speed (fast or slow), pathology itself and disease stage may impact on the raw signal (Figure 3, (2)) and therefore influence algorithm performance. In addition other technical considerations need to be taken into account. Many commercial devices adopt black box designs with un-validated firmware/software [79] which account for at least some of the significant disagreements in reported results [80, 81]. Other uncertainties due to externally induced motion (e.g. cars, lifts) also impact on accuracy to detect features of interest [81]. Static and dynamic re-calibration of WTCD to account for possible axis misalignment or sensor alterations due to damage (device dropped, contact with water etc.) is also advised [82], however rarely undertaken because procedures are complicated and expensive. Further sources of variability are also introduced through changes in external factors such as weather, mood

[1] https://osf.io/2wuv9/

or medication, influencing analysis of the signal. Collectively these result in errors and decreased confidence in outcomes and conformity to everyday use. Algorithm development will ultimately refine extraction and a joint approach such as use of secondary data sources will aid interpretation, for example data from patients' diaries, testimony from carers, and use of clinical records [83]. All of these potential sources of error should be considered and some suggestions are provided in Table 2.

Determining optimal outcome measures. Table 1 shows the vast range of outcomes reported. Standardised measurement is urgently needed with a clear rationale for selection of outcomes from which clinimetric testing will allow a refined battery of measures to emerge to encourage harmonisation across studies. Examples of measurement frameworks have been described [46, 49] including our own *micro* and *macro* level structure used throughout this paper [47]. Others [37, 38, 45, 57, 61] beside volume outcomes (e.g. total number of walking bouts, etc.) defined as '*quantity*' metrics, use novel frequency-based outcomes to characterise gait (a) symmetry, variability and stability (e.g. harmonic ratio, amplitude of dominant frequency, dynamic stability, etc.) defined broadly as '*quality'* metrics. These novel *quality* measures, although very promising for discriminative validity, may be difficult to interpret in clinical practice.

<Figure 2>

<Figure 3>

Free-living monitoring in PD: where to next?

Modern devices incorporate a range of inertial sensors such as accelerometers, gyroscopes, magnetometers with Bluetooth connectivity which constitute cutting edge WTCD. While use is currently limited to controlled settings, improvements in battery technology will improve the accuracy of measurement addressing some of the challenges highlighted earlier. Moreover, novel methods for advanced data processing are being developed to reduce computational load with advanced computational processing carried out remotely via smartphone or in the cloud extending the application of WTCD [84]. Studies have also investigated the use of smart phones (and audio devices)

which regularly come with the necessary hardware to quantify symptoms, movement or gait [85]. These devices capture, analyse and relay information via cellular or other wireless networks and also provide a more comprehensive assessment such as the addition of a microphone for use with speech analysis algorithms in PD diagnosis [86, 87] and visual displays to facilitate applications (apps) for the study of cognition [88]. Rigorous device testing however is needed to ensure confidence in their application.

Long term monitoring via a smart phone facilitates network interconnectivity and integration to the Internet of Things (IoT) [5], through delayed or real-time uploading of data to cloud computing infrastructures. Data can be relayed to the patient (bio-feedback) via unobtrusive displays, haptic and audible cues. Data is stored and sent to clinicians for tracking disease progression, optimising disease management and providing further, more clinically informed feedback to the patient. Data storage and data access on this scale constitutes 'big data analytics'. Developments in this field can expand assessment to capture the 'lived experience' or 'lifespace' of PD, capturing the extent to which people travel and their patterns of movement within the community [89]. This is exemplified by a recent collaborative project between the Michael .J Fox Foundation and Apple utilising their projects, FoxInsight[2] and the Apple ResearchKit[3] (inc. the Parkinson mPower app[4] available via iTunes), respectively.

Collection of data on the scale and in a free-living context raises new ethical challenges with respect to acquisition, analysis and storage. Current ethical reviews may not be sufficient to identify modern issues [90]. Technology and terminology has evolved faster than legal and ethical systems and unforeseen issues can emerge [91]. Informed, principled, and collaborative experimentation are therefore necessary to ensure privacy and confidentiality, and compliance with ethical principles.

Conclusions and recommendations

There is no doubting the possibilities and potential of real world monitoring and assessment of clinical features for people with PD. It is conceivable to imagine a future where *micro* level data is used to enhance diagnostics, measure efficacy of intervention and monitor disease progression, and

[2] The Michael J. Fox Foundation for Parkinson's Research, https://foxinsight.michaeljfox.org/
[3] Apple Inc., http://www.apple.com/uk/researchkit/
[4] http://parkinsonmpower.org/

predict risk of disease, falls and cognitive decline. *Macro* level data, on the other hand, reflects the global burden of disease and impact of therapy. Both sources of data provide insights into personalised treatment. As this special issue in the journal indicates, this is a rapidly developing field. However, much work remains before widespread clinical adoption is a reality. We highlight key recommendations and some practical solutions to move this field forward (Table 2). These challenges are likely to be met most effectively by adopting a multidisciplinary approach between key stakeholders such as clinicians, patients, engineers, computer scientists, and statisticians.

Acknowledgments

The authors would like to acknowledge Dr. Brook Galna for his support in editing Figure 1.

Authors' roles

SDD: Manuscript organisation, writing, review and critique.

AG: Manuscript writing, review and critique.

CM: Manuscript writing, review and critique.

SL: Manuscript writing, review and critique.

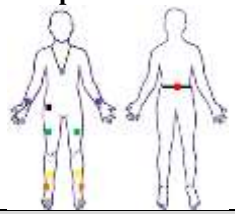
LR: Manuscript conception, writing, review and critique.

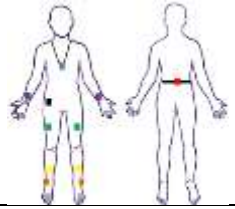
Financial Disclosures of all authors

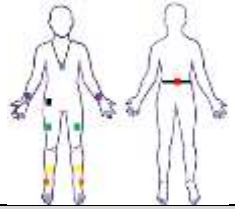
SDD is supported by the V-Time project, which is a European Union 7th Framework Programme (FP7) under the Health theme (FP7 - 278169). AG, SL and LR are supported by the National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre and Unit based at Newcastle upon Tyne Hospitals NHS Foundation Trust and Newcastle University. CM is supported by the EPSRC Frontier Engineering Awards, Grant Reference No. EP/K03877X/1 and by the MRC and Arthritis Research UK as part of the MRC – Arthritis Research UK Centre for Integrated research into Musculoskeletal Ageing (CIMA). The views expressed are those of the authors and not necessarily those of the NHS or NIHR or the Department of Health.

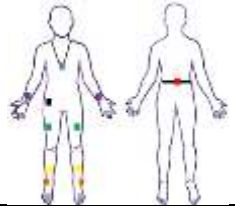
Tables

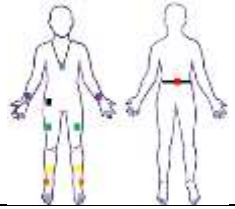
Table 1: Studies examining free-living monitoring of Parkinson's disease (PD) using wearable technology and connected devices (WTCD).
Number and position of WTDC used in each study is detailed in column two using a colour code (blue = chest, violet = wrist, black = pocket, green = thigh, yellow = shank, orange = ankle, grey = foot, red = lower back).
† Clinical feature/ activity detected or measures has been classified using three types of validity: 1) accurate detection of clinical feature/ method of appraisal: the ability of WCTD algorithms to accurately detect a clinical feature/activity which is comparable to detection by another means - in the study cited or previous studies (e.g. self-report, EMG); 2) criterion validity: the association between WTCD-derived outcomes and measures such as clinical scales; and 3) discriminative validity: the ability of WTCD-derived outcomes to discriminative between groups. Formal testing of utility (feasibility/compliance intentionally tested and reported) of WTCD is also reported.

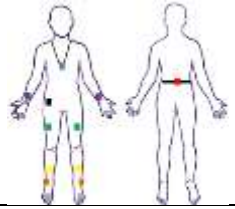
Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
Motor symptoms and medication response							
Das et al. (2012), 2 PD, 4*[21]	Accelerometers --■-- ■■■■	Dyskinesia, tremor	Yes, against patients' diaries using weakly supervised machine learning technique.	Acceleration derived features (Mean energy, high frequency energy content, correlation, frequency domain entropy)	No	No	No
Griffiths et al. (2012), 34 PD/10 OA, 10[18]	Parkinson's KinetiGraph (PKG; Global Kinetics Corporation) ■	Bradykinesia, dyskinesia	Yes, for bradykinesia against dot slide task measure (specificity 88%, sensitivity 95%) during scripted tests.	Acceleration derived features: Mean Spectral Power within specific bands, peak, amount of time with no movement	Yes, dyskinesia against the AIMS score and both dyskinesia and bradykinesia against UPDRS III and IV	No	No
Mera et al. (2012), 10 PD/ 10 OA, 3-6[19]	Kinesia™ ■--	Motor tasks, tremor, bradykinesia, motor fluctuations	No	Symptoms severity scale (0-4 points), voluntary movement threshold evaluated with gyroscope derived features (RMS, peak of power spectrum)	Yes, for tremor and bradykinesia. Potential issues of recognition when the 2 symptoms overlap. Yes against videos in the lab for symptom severity scale validated against UPDRS tremor and MBRS speed, amplitude and rhythm scores in previous work [75, 92]	No	Yes, formal testing previous work [75]
Pastorino et al. (2013), 2 PD, 7	ALA-6g (PERFORM) --■-- ■■■■	Akinesia, ON/OFF state	Yes, 'proof of concept' validation	Level of akinesia	No	No	Yes, formal testing

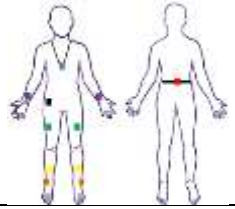
Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
(but 32 hours analysed) [17]			against patients' diaries				
Tzallas et al. (2014), 12 PD, 5 (8 hours per day)[22]	ALA-6g (PERFORM) --■-- ■■■■	Tremor, LID, Bradykinesia, FOG	Yes, in the lab and during structured test (e.g. for FOG events Opening door/ Straight 10m walking) against video annotations.	Acceleration derived measures (time and frequency domains, RMS, range, entropy, energy)	Yes, machine learning and leave one out validation technique validated in the lab and applied in free-living conditions and compared against patients' diaries. Use of videos in the lab for assessing symptoms severity using UPDRS.	No	Yes, formal testing
Ferreira et al. (2015), 11 PD, 12 weeks [23]	SENSE-PARK System ■■■	Gait, hypokinesia, dyskinesia, tremor, sleep	No/NA (feasibility study and usability)	NA	NA	No	Yes, formal testing
Hammerla et al. (2015), 34 PD, 7[24]	Axivity AX3 ■■	Sleeping, ON/OFF state, dyskinesia	Yes, in the lab (against video recordings) using machine learning and leave one out validation technique, in free-living conditions results are compared against patients' diaries. Model pre-trained in free-living conditions did not give good results (laboratory data is a poor model for naturalistic behaviours)	Acceleration derived measures (magnitude, jerk, power spectral density, etc.)	No	No	Yes, formal testing but in subsequent work [93]
Horne et al. (2015), 64 PD/38 OA, 10[20]	Parkinson's KinetiGraph (PKG; Global Kinetics Corporation) ■	Bradykinesia, dyskinesia, fluctuations	Yes, against measures of bradykinesia and dyskinesia (previous work see Griffiths 2012)	Fluctuation Score based on Interquartile Range of bradykinesia and dyskinesia scores.	Yes, against clinical scores derived measure	Yes	No

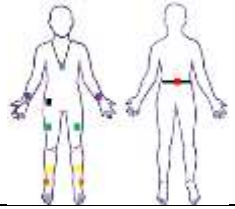
Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
Sleep							
Prudon et al. (2013), 106 PD/99 OA, 3 nights[94]	Acti-watch, Camntech	Leg movements during sleep	Yes, in patients with periodic leg movement (against electromyography), previous work	Periodic leg movements index	Yes, against disease severity	No	No
Louter et al. (2015), 11 PD, 2 nights[25]	Dynaport McRoberts	Turning during sleep	Yes, against polysomnography in adults with obstructive sleep apnoea syndrome, previous work [95]	Acceleration derived measures (e.g. mean) and axial movement measures (frequency, size, duration, speed)	Yes, against Acti-watch but in young healthy adults previous work [95]	Yes	Yes, no formal testing, previous work
Sringean et al. (2015), 19 PD, 1 night[26]	NIGHT-Recorder system	Turning, Standing	No, video and sleep diaries collected but validity not formally tested.	Acceleration and gyroscope derived measures (duration of sleep, axial movements, velocity, etc.)	Yes, against clinical scores (UPDRS axial score, item #28, etc.)	Yes	Yes, no formal testing, no adverse events reported
Falls and Falls Risk							
Weiss et al. (2013), 71 OA, 3[35]	Dynaport McRoberts	Walking (at least 60s)	No	Number of walking bouts, walking duration, total number of steps, median number of steps per bout, bout duration, cadence, step and stride regularity, frequency domain measures (harmonic ratio, amplitude, slope and width of dominant frequency), step duration, step symmetry, acceleration range, etc.	Yes, against clinical scores of fall risk and laboratory based measures	Yes	No
Weiss et al. (2014), 107 PD, 3[36]	Dynaport McRoberts	Walking (at least 60s)	No	Number of walking bouts, % of activity duration, total number of steps, median number of steps per bout, bout duration, cadence, stride regularity, frequency domain measures (harmonic ratio,	Yes, against clinical scores of fall risk	Yes	Yes, no formal testing, data loss reported.

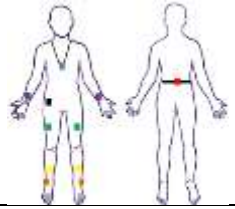
Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
				amplitude and width of dominant frequency), etc.			
Brodie et al. (2015), 18 EF, 58 (average)[40]	Senior Mobility Monitor (SMM, Philips) --▮--	Walking (at least 3 or 8 steps)	No	Steps per day, walking bouts per day, steps per bout, cadence, distribution of bout length	No	Yes	No
Hiorth et al. (2015), 48 PD, 7[41]	activPAL ▮	Sedentary behaviour/ standing/ walking	Yes, but not formal in PD. Previous work in OA against other accelerometer [96] and video recordings in people with rheumatoid arthritis during simulation of ADL in the laboratory [97]	Volume (e.g. total number of sedentary/standing/walking bouts), pattern (α), variability of sedentary bouts and number of strides per walking bout.	Yes, against clinical scores	Yes	No
Mactier et al. (2015), 111 PD, 7[39]	activPAL ▮	Walking	Yes, but not formal in PD. Previous work in OA against other accelerometer [96] and video recordings in people with rheumatoid arthritis during simulation of ADL in the laboratory [97]	Volume (e.g. total number of walking bouts), pattern (α), variability of bouts, accumulation of stepping bouts	No	Yes	No
Rispens et al. (2015), 113 OA, 14[38]	Dynaport McRoberts ▮	Walking (at least 10s)	Yes, previous work in OA [98] for walking volume parameters against videos, no for gait characteristics.	Acceleration based outcomes: gait speed, speed variability, stride time, stride regularity, stride time variability, stride frequency, frequency domain measures (harmonic ratio, amplitude, slope and width of dominant frequency), etc.	Yes, measures against self-reported fall history	No	No
van Schooten et al. (2015), 169 OA, 8[37]	Dynaport McRoberts ▮	Walking (at least 10s), sitting, lying, and standing	Yes, previous work in OA [98] for walking volume parameters against videos, no for	Total duration of walking, sitting, standing, and lying per day, number of	Yes, against falls history	Yes	No

Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
			gait characteristics.	strides, number of walking bouts, duration of bouts, number of transitions. Gait characteristics: gait speed, stride frequency, stride length frequency domain measures (harmonic ratio, power at dominant frequency), etc.			
Kangas et al. (2015), 16 OA, 5-155[32]	CareTech Ab -- --	Falls‡	Yes, fall event against care personnel's reports and in previous work in OA during simulation of fall events in controlled conditions [99] in OA	Fall event with alarm generation	No	No	Yes, based on alarm accuracy
Freezing of Gait (FOG)							
Moore et al. (2013), 25 PD, NA[43]	Xsens MTx	Turning/ walking (TUG)‡	Yes, in the laboratory for FOG event against video recordings	FOG event through acceleration derived frequency measures (power spectrum, etc.).	No	No	No
Tripoliti et al. (2013), 11 PD/5 OA, NA[44]	Body Sensor AGYRO, AGYRO links, ANCO S.A.	Walking, FOG detection‡	Yes, against video recordings and visual inspection during structured test (Opening door/ Straight 10m walking) sing different classification algorithms and cross-validations	FOG detection through entropy of WTCD signal	No	No	No
Weiss et al. (2015), 72 PD, 3[45]	Dynaport McRoberts	Walking (at least 60s)	No	Number of walking bouts, % of activity duration, total number of steps, median number of steps per bout, bout duration, cadence, stride regularity, frequency domain	Yes, against clinical scores (FOG questionnaire)	Yes	No

Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
				measures (harmonic ratio, width of dominant frequency), etc.			
			Gait				
Cancela et al. (2011), 10 PD, 1 (not clear)[58]	ALA-6g (PERFORM) --■-- ■■■■	Walking (on vs off medication)	Yes, only for step frequency during 10m scripted protocol against visual examination	Step frequency, stride length and speed, entropy, arm swing	No	Yes, only for entropy in previous work [100]	No
Weiss et al. (2011), 22 PD/17 OA (1PD/1CL at home), 3[57]	Mobi8 ■	Walking (during scripted test in the lab and during simulation of ADL and free-living)	No	Acceleration derived measures (time and frequency domains): stride time, stride time variability, amplitude, width, slope of dominant frequency, etc.	Yes, against clinical scores	Yes	No
Cancela et al. (2014), 11 PD, 5-7 (8 hours per day)[59]	ALA-6g (PERFORM) --■-- ■■■■	Walking	Yes, only for step frequency, previous work (see Cancela 2011)	Step frequency, step velocity, stride length, entropy	No	Yes, only for entropy in previous work [100]	Yes, formal testing and also assessed in separate study [76]
Herman et al. (2014), 110 PD, 3[61]	Dynaport McRoberts ■	Walking (at least 60s)	No	Total number of activity bouts, total % of activity duration, total number of steps, mean activity bout duration, median number of steps per bout, cadence, stride regularity, amplitude of dominant frequency, width of dominant frequency, stride regularity, harmonic ratio, Phase Coordination Index.	Yes, previous work	Yes, previous work	No.
Weiss et al. (2015), 107 PD, 3[60]	Dynaport McRoberts ■	Walking (at least 60s)	No	Total % of activity duration, total number of steps, cadence, amplitude of dominant frequency, stride regularity, harmonic ratio,	Yes, previous work	Yes, previous work	Yes, no formal testing, data loss reported

Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
				Phase Coordination Index.			
Del Din et al. (2016), 47 PD/50 OA, 7 [49]	Axivity AX3	Walking (at least 3 steps)	No	14 gait characteristics: mean step time, stance time, swing time, step length, step velocity, step time variability, stance time variability, swing time variability, step length variability, step velocity variability, step time asymmetry, stance time asymmetry, swing time asymmetry, step length asymmetry.	Yes, gait characteristics validated against laboratory reference (previous work [53])	Yes	No
				Timed-up-and-go (TUG)			
Zampieri et al. (2011), 6 PD/8 OA, 1[62]	Physilog	Walking/turning/postural transitions ǂ	Yes, in previous work[101]	Cadence, stride velocity, stride length, peak arm velocity, turning velocity	No	Yes	No
Smith et al. (2016), 12 OA, 5[63]	SHIMMER	Walking/turning ǂ	No	Time to complete test, cadence, gait characteristics (step time, stride time, stride length, stride velocity, etc.), turning magnitude, etc.	No	No	No
				Turning			
El-Gohary et al. (2013), 12 PD/18 OA, 7*[65]	Opal(ADPM) -- -- in the lab / Opal(ADPM) -- -- at home	Turning/ walking (at least 10s)	Yes, in the lab against motion analysis system and video recordings	Number of turns, peak velocity, mean velocity, duration of turn	No	Yes	No
Mancini et al. (2015), 13 PD/8 OA, 7*[64]	Opal(ADPM) -- --	Turning/ walking (at least 10s)	Yes, in the lab (previous work, see El-Gohary 2013)	Number of turns/hour, turn angle, turn duration, number of steps/turn, turn mean velocity and coefficient of variation of these measures.	Yes	Yes	Yes, no formal testing, report of 'ease' of use.
				Ambulatory activity and sedentary behaviour			
Chastin et al. (2010), 17 PD/17 OA, 7[71]	activPAL	Sedentary behaviour	Yes, but not formal in PD. Previous work in OA against other	Volume of sedentary bouts, pattern (α), pattern of accumulation of bouts (GINI	No	Yes	No

Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
			accelerometer [96] and video recordings in people with rheumatoid arthritis during simulation of ADL in the laboratory [97]	index)			
Dontje et al. (2013), 467 PD, 14[70]	TracmorD, Philips/ --■-- or --■-- or ■	Physical Activity/Sedentary behaviour	Yes, against doubly labeled water technique (correlation) in adults but not in PD [102]	Energy expenditure, time spent in activities, distribution of activities, etc.	Yes	No	No
Benka Wallen et al. (2015), 95 PD, 7[103]	ActiGraph GT3X+ -- ■–	Physical Activity/Sedentary behaviour/ Steps (60s epochs)	Yes, in young adults under controlled conditions by visual observation but not in PD [104]	Volume (magnitude vector of acceleration) and time spent in physical activities, steps per day, etc.	No	No	No
Lim et al. (2010), 153 PD, 1[74]	Vitaport3, TEMEC Instruments BV ■■ ■■	Sitting, standing, walking	Yes in PD against video (under controlled conditions), previous work [105]	% of time spent on dynamic, static, sitting, standing or walking activities, number of walking bouts > 5s and > 10s	No	No	No
Cavanaugh et al. (2012), 33 PD, 7[72]	StepWatch 3 Step Activity Monitor (SAM) ■	Walking (average every 60s)	Yes, for stride count in PD against instrumented walkway in the lab, previous work [106]	Total number of steps, maximum output for steps, number of minutes with > 100 steps, number and duration of walking bouts, peak activity index, % of day spent inactive	No	No	Yes, not formal testing, reasons for data loss and attrition in sensor acceptability after 1 year with decrease in participant use reported
Rochester et al. (2012), 17 PD, 7[68]	activPAL ■	Walking	Yes, but not formal in PD. Previous work in OA against other accelerometer [96] and video recordings in	Volume of walking bouts, pattern of accumulation of bouts (GINI index) and diversity of bouts, distribution and variability of bouts (S_2)	Yes	No	No

Study (Year), N, Length of recording	WTCD and placement 	Clinical feature/ Activity	Accurate detection of clinical feature: Method of appraisal†	Measures	Criterion Validity†	Discriminative Validity†	Utility
			people with rheumatoid arthritis during simulation of ADL in the laboratory [97]				
Lord et al. (2013), 89 PD/97 OA, 7[66]	activPAL	Walking	Yes, but not formal in PD. Previous work in OA against other accelerometer [96] and video recordings in people with rheumatoid arthritis during simulation of ADL in the laboratory [97]	Volume of walking bouts, pattern (α), time spent walking in short-medium or long bouts, frequency and variability of bouts (S_2)	Yes	Yes	No
Cavanaugh et al. (2015), 17 PD, 7[73]	StepWatch 3 Step Activity Monitor (SAM)	Walking (average every 60s)	Yes, for stride count in PD against instrumented walkway in the lab, previous work (see Cavanaugh 2012)	Mean daily steps, maximum output for steps, Moderate intensity minutes (number of minutes with > 100 steps)	Yes	No	Yes, not formal testing, reasons for data loss and attrition in sensor acceptability after 2 years with decrease in participant use reported

ADL = Activities of Daily Living; Alpha = α; Lab = Laboratory; Length of recording= number of weeks/days/minutes of recording; MBRS = Modified Bradykinesia Rating Scale; min = miutes; N = number of participants; OA = Older Adults; PD = Parkinson's disease; RMS = Root Mean Square; UPDRS = Unified Parkinson's Disease Rating Scale; % = Percentage; *Night excluded; ǂ = scripted protocol/supervised conditions used.

Table 2: Practical solutions and broad recommendations for WTCD-related research challenges.

Recommendation	Practical solutions
Adopt standardised definition of activity/clinical feature	• Justify definition of activity/clinical feature with respect to earlier work & clinical expertise. • Adopt interdisciplinary collaboration for optimal process, choice of equipment, protocol, data processing and outcomes adhering to research question(s).
Select equipment depending on research/clinical question; evaluate trade-off between information needed & equipment available.	• Consider optimal technical specifications (e.g. sampling frequency, type of data collected; battery life) for outcome measures. • Use WTCD with established utility, acceptability and cost-effectiveness, otherwise plan to include tests of utility and acceptability as part of the study. • Ensure transparency of all aspects of technology used (specifications, data collection, data pre-processing).
Use standardised protocols & validation procedures for algorithms for comparability & reproducibility across studies (e.g. accurate detection of activity/clinical feature, criterion & discriminative validity).	• Justify use of standardised protocol & methods to define activities/clinical features. • Use algorithms previously validated four the current application or provide validation results for novel algorithms. • Use appropriate gold standards (e.g. video recording) to validate outcomes/metrics in free-living conditions, not limiting validation to scripted protocols or controlled conditions. • Account for influence of context and disease severity on algorithm performance. • If proprietary software is used ensure transparency of manufacturer algorithms or report published validated algorithm.
Achieve consensus for summary outcomes for comparability across studies.	• Use WTCD-based outcomes validated in free-living; or provide validation results in the current study using semi-structured activities. • Describe (if any) dependence of chosen summary outcomes & on chosen data processing/algorithm.

Figures

Figure 1:
Use of wearable technology and connected devices (WTCD) (adapted with permission from previous work) [47] A) *macro* level quantification of activities over an extended period of time (volume, patterns and variability); (B) bouts of activities (e.g. lying (sleeping), walking, sitting); (C-H) *micro* level quantification from specific events: C) and D) postural transitions, E) shuffling, F) gait, G) turning, H) freezing of gait (FOG) and fall.

Figure 2:
Examples of linear and non-linear approaches to activity data analysis: volume and pattern metrics for two subjects (Subject 1 and 2) (published with permission) [68].
A1 and A2 - Patterns of activity indicating bouts of sedentary, standing and walking at different stepping rates (cadences).
B1 and B2 - Volume Metrics: total walking time for the two subjects is equal but made up of walking bouts at different cadences.
C - Pattern Metrics: (i) and (ii) distribution of walking bouts for these two subjects with equal mean (M) and different dispersion (S2). C (iii) Accumulation pattern of walking time for subject 1 and 2; subject 2 tends to accumulate walking time with predominantly longer periods.

Figure 3:
Challenges/limitations of free-living measurement using examples from gait in free-living collected with a single accelerometer-based WTCD. Data (unpublished) from the Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation-GAIT (ICICLE-GAIT) study [107].

Panel (1) – Definition of feature of interest (e.g. walking):

A) Impact of "selected" definition of walking on data processing: different threshold of walking bout length and (ghost) maximum resting period (MRP) between consecutive walking bouts can be utilised.
   Examples: (i): use of walking bout threshold of 60s and no MRP (MRP = 0s) (only bouts longer than 60 s will be considered); (ii): use of walking bout threshold of 3 steps and no MRP (MRP = 0s); (iii) use of walking bout threshold of 3 steps and MRP = 5s.
B) Impact of choice in A) on *macro* outcomes (e.g. number of bouts considered, total number of steps reported for people with Parkinson's disease (PD) and controls (CL)). For example using definition (i) only a small percentage of all the walking bouts will be considered (bouts > 60s only) and therefore fewer steps will be reported if compared to results of using definition (ii).
C) Impact of choice in A) on *micro* gait characteristics (e.g. reported step velocity may vary across studies due to choice of definition ((i), (ii) or (iii)).

Panel (2) – Influence of free-living protocol on data:

Walking speed changes with respect to the environment, task, and disease severity which influences the accelerometer raw signal (D) impacting on algorithm performance and evaluation of outcomes (E).

References

1. Lowe SA, Olaighin G. Monitoring human health behaviour in one's living environment: a technological review. Med Eng Phys 2014;36(2):147-168.
2. Robles-Garcia V, Corral-Bergantinos Y, Espinosa N, et al. Spatiotemporal Gait Patterns During Overt and Covert Evaluation in Patients With Parkinson s Disease and Healthy Subjects: Is There a Hawthorne Effect? Journal of applied biomechanics 2015;31(3):189-194.
3. Maetzler W, Rochester L. Body-worn sensors-the brave new world of clinical measurement? Mov Disord 2015;30(9).
4. Steins D, Dawes H, Esser P, Collett J. Wearable accelerometry-based technology capable of assessing functional activities in neurological populations in community settings: a systematic review. J Neuroeng Rehabil 2014;11:36.
5. Pasluosta C, Gassner H, Winkler J, Klucken J, Eskofier B. An Emerging Era in the Management of Parkinson's disease: Wearable Technologies and the Internet of Things. IEEE J Biomed Health Inform 2015;19(6):1873-1881.
6. Awais M, Mellone S, Chiari L. Physical activity classification meets daily life: Review on existing methodologies and open challenges. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2015;2015:5050-5053.
7. Hobert MA, Maetzler W, Aminian K, Chiari L. Technical and clinical view on ambulatory assessment in Parkinson's disease. Acta neurologica Scandinavica 2014;130(3):139-147.
8. Maetzler W, Domingos J, Srulijes K, Ferreira JJ, Bloem BR. Quantitative wearable sensors for objective assessment of Parkinson's disease. Mov Disord 2013;28(12):1628-1637.
9. Godinho C, Domingos J, Cunha G, et al. A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. J Neuroeng Rehabil 2016;13(1):24.
10. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 2008;23(15):2129-2170.
11. Ossig C, Antonini A, Buhmann C, et al. Wearable sensor-based objective assessment of motor symptoms in Parkinson's disease. J Neural Transm (Vienna) 2016;123(1):57-64.
12. Papapetropoulos S, Mitsi G, Espay AJ. Digital Health Revolution: Is it Time for Affordable Remote Monitoring for Parkinson's Disease? Frontiers in neurology 2015;6:34.
13. Godfrey A, Lara J, Del Din S, et al. iCap: Instrumented assessment of physical capability. Maturitas 2015;82(1):116-122.
14. Schoneburg B, Mancini M, Horak F, Nutt JG. Framework for understanding balance dysfunction in Parkinson's disease. Mov Disord 2013;28(11):1474-1482.
15. Horak F, King L, Mancini M. Role of body-worn movement monitor technology for balance and gait rehabilitation. Phys Ther 2015;95(3):461-470.
16. Cancela J, Pansera M, Pastorino M, Pastor L, Arredondo MT. Automatic assessment of bradykinesia severity in patients with Parkinson’s disease. 7th International Conference on Wearable Micro and Nano Technologies for Personalized Health 2010;7th International Conference on Wearable Micro and Nano Technologies for Personalized Health.
17. Pastorino M, Cancela J, Arredondo MT, Pastor-Sanz L, Contardi S, Valzania F. Preliminary results of ON/OFF detection using an integrated system for Parkinson's disease monitoring. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2013;2013:941-944.

18. Griffiths RI, Kotschet K, Arfon S, et al. Automated assessment of bradykinesia and dyskinesia in Parkinson's disease. Journal of Parkinson's disease 2012;2(1):47-55.
19. Mera TO, Heldman DA, Espay AJ, Payne M, Giuffrida JP. Feasibility of home-based automated Parkinson's disease motor assessment. J Neurosci Methods 2012;203(1):152-156.
20. Horne MK, McGregor S, Bergquist F. An objective fluctuation score for Parkinson's disease. PloS one 2015;10(4):e0124522.
21. Das S, Amoedo B, De la Torre F, Hodgins J. Detecting Parkinsons' symptoms in uncontrolled home environments: a multiple instance learning approach. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2012;2012:3688-3691.
22. Tzallas AT, Tsipouras MG, Rigas G, et al. PERFORM: a system for monitoring, assessment and management of patients with Parkinson's disease. Sensors (Basel, Switzerland) 2014;14(11):21329-21357.
23. Ferreira JJ, Godinho C, Santos AT, et al. Quantitative home-based assessment of Parkinson's symptoms: the SENSE-PARK feasibility and usability study. BMC neurology 2015;15:89.
24. Hammerla NY, Fisher JM, Andras P, Rochester L, Walker R, Plötz T. PD Disease State Assessment in Naturalistic Environments using Deep Learning. Twenty-Ninth AAAI Conference on Artificial Intelligence; 2015.
25. Louter M, Maetzler W, Prinzen J, et al. Accelerometer-based quantitative analysis of axial nocturnal movements differentiates patients with Parkinson's disease, but not high-risk individuals, from controls. Journal of neurology, neurosurgery, and psychiatry 2015;86(1):32-37.
26. Sringean J, Taechalertpaisarn P, Thanawattano C, Bhidayasiri R. How well do Parkinson's disease patients turn in bed? Quantitative analysis of nocturnal hypokinesia using multisite wearable inertial sensors. Parkinsonism Relat Disord 2015;23:10-16.
27. Bourke AK, O'Donovan KJ, Nelson J, GM OL. Fall-detection through vertical velocity thresholding using a tri-axial accelerometer characterized using an optical motion-capture system. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2008;2008:2832-2835.
28. Bourke AK, Torrent M, Parra X, Catala A, Nelson J. Fall algorithm development using kinematic parameters measured from simulated falls performed in a quasi-realistic environment using accelerometry. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2011;2011:4449-4452.
29. Bourke AK, van de Ven P, Gamble M, et al. Assessment of waist-worn tri-axial accelerometer based fall-detection algorithms using continuous unsupervised activities. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2010;2010:2782-2785.
30. Kangas M, Konttila A, Lindgren P, Winblad I, Jamsa T. Comparison of low-complexity fall detection algorithms for body attached accelerometers. Gait Posture 2008;28(2):285-291.
31. Kangas M, Konttila A, Winblad I, Jamsa T. Determination of simple thresholds for accelerometry-based parameters for fall detection. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2007;2007:1367-1370.

32. Kangas M, Korpelainen R, Vikman I, Nyberg L, Jamsa T. Sensitivity and false alarm rate of a fall sensor in long-term fall detection in the elderly. Gerontology 2015;61(1):61-68.
33. Henderson EJ, Lord SR, Close JC, Lawrence AD, Whone A, Ben-Shlomo Y. The ReSPonD trial--rivastigmine to stabilise gait in Parkinson's disease a phase II, randomised, double blind, placebo controlled trial to evaluate the effect of rivastigmine on gait in patients with Parkinson's disease who have fallen. BMC neurology 2013;13:188.
34. Markle-Reid M, Browne G, Gafni A, et al. A cross-sectional study of the prevalence, correlates, and costs of falls in older home care clients 'at risk' for falling. Canadian journal on aging = La revue canadienne du vieillissement 2010;29(1):119-137.
35. Weiss A, Brozgol M, Dorfman M, et al. Does the evaluation of gait quality during daily life provide insight into fall risk? A novel approach using 3-day accelerometer recordings. Neurorehabil Neural Repair 2013;27(8):742-752.
36. Weiss A, Herman T, Giladi N, Hausdorff JM. Objective assessment of fall risk in Parkinson's disease using a body-fixed sensor worn for 3 days. PloS one 2014;9(5):e96675.
37. van Schooten KS, Pijnappels M, Rispens SM, Elders PJ, Lips P, van Dieen JH. Ambulatory fall-risk assessment: amount and quality of daily-life gait predict falls in older adults. J Gerontol A Biol Sci Med Sci 2015;70(5):608-615.
38. Rispens SM, van Schooten KS, Pijnappels M, Daffertshofer A, Beek PJ, van Dieen JH. Identification of fall risk predictors in daily life measurements: gait characteristics' reliability and association with self-reported fall history. Neurorehabil Neural Repair 2015;29(1):54-61.
39. Mactier K, Lord S, Godfrey A, Burn D, Rochester L. The relationship between real world ambulatory activity and falls in incident Parkinson's disease: influence of classification scheme. Parkinsonism Relat Disord 2015;21(3):236-242.
40. Brodie M, Lord S, Coppens M, Annegarn J, Delbaere K. Eight weeks remote monitoring using a freely worn device reveals unstable gait patterns in older fallers. IEEE transactions on bio-medical engineering 2015;62(11):2588-2594.
41. Hiorth YH, Larsen JP, Lode K, et al. Impact of falls on physical activity in people with Parkinson's disease. Journal of Parkinson's disease 2015.
42. Snijders AH, Haaxma CA, Hagen YJ, Munneke M, Bloem BR. Freezer or non-freezer: clinical assessment of freezing of gait. Parkinsonism Relat Disord 2012;18(2):149-154.
43. Moore ST, Yungher DA, Morris TR, et al. Autonomous identification of freezing of gait in Parkinson's disease from lower-body segmental accelerometry. J Neuroeng Rehabil 2013;10:19.
44. Tripoliti EE, Tzallas AT, Tsipouras MG, et al. Automatic detection of freezing of gait events in patients with Parkinson's disease. Computer methods and programs in biomedicine 2013;110(1):12-26.
45. Weiss A, Herman T, Giladi N, Hausdorff JM. New evidence for gait abnormalities among Parkinson's disease patients who suffer from freezing of gait: insights using a body-fixed sensor worn for 3 days. Journal of neural transmission (Vienna, Austria : 1996) 2015;122(3):403-410.
46. Lord S, Galna B, Verghese J, Coleman S, Burn D, Rochester L. Independent domains of gait in older adults and associated motor and nonmotor attributes: validation of a factor analysis approach. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 2013;68(7):820-827.
47. Lord S, Galna B, Rochester L. Moving forward on gait measurement: toward a more refined approach. Mov Disord 2013;28(11):1534-1543.
48. Mollenhauer B, Rochester L, Chen-Plotkin A, Brooks D. What can biomarkers tell us about cognition in Parkinson's disease? Mov Disord 2014;29(5):622-633.

49. Del Din S, Godfrey A, Galna B, Lord S, Rochester L. Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. Journal of NeuroEngineering and Rehabilitation 2016;In Press.
50. Salarian A, Russmann H, Vingerhoets FJ, et al. Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring. IEEE transactions on bio-medical engineering 2004;51(8):1434-1443.
51. Esser P, Dawes H, Collett J, Feltham MG, Howells K. Validity and inter-rater reliability of inertial gait measurements in Parkinson's disease: A pilot study. Journal of neuroscience methods 2012;205(1):177-181.
52. Esser P, Dawes H, Collett J, Feltham MG, Howells K. Assessment of spatio-temporal gait parameters using inertial measurement units in neurological populations. Gait Posture 2011;34(4):558-560.
53. Del Din S, Godfrey A, Rochester L. Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and Parkinson's disease: toward clinical and at home use. IEEE J Biomed Health Inform 2016;20(3):838-847.
54. Trojaniello D, Cereatti A, Pelosin E, et al. Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait. J Neuroeng Rehabil 2014;11:152.
55. Trojaniello D, Ravaschio A, Hausdorff JM, Cereatti A. Comparative assessment of different methods for the estimation of gait temporal parameters using a single inertial sensor: application to elderly, post-stroke, Parkinson's disease and Huntington's disease subjects. Gait Posture 2015;42(3):310-316.
56. Brodie MA, Coppens MJ, Lord SR, et al. Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different. Medical & biological engineering & computing 2015;54(4):663-674.
57. Weiss A, Sharifi S, Plotnik M, van Vugt JP, Giladi N, Hausdorff JM. Toward automated, at-home assessment of mobility among patients with Parkinson disease, using a body-worn accelerometer. Neurorehabil Neural Repair 2011;25(9):810-818.
58. Cancela J, Pastorino M, Arredondo MT, et al. Gait assessment in Parkinson's disease patients through a network of wearable accelerometers in unsupervised environments. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2011;2011:2233-2236.
59. Cancela J, Pastorino M, Arredondo MT, Nikita KS, Villagra F, Pastor MA. Feasibility study of a wearable system based on a wireless body area network for gait assessment in Parkinson's disease patients. Sensors (Basel, Switzerland) 2014;14(3):4618-4633.
60. Weiss A, Herman T, Giladi N, Hausdorff JM. Association between Community Ambulation Walking Patterns and Cognitive Function in Patients with Parkinson's Disease: Further Insights into Motor-Cognitive Links. Parkinson's disease 2015;2015:547065.
61. Herman T, Weiss A, Brozgol M, Giladi N, Hausdorff JM. Gait and balance in Parkinson's disease subtypes: objective measures and classification considerations. J Neurol 2014;261(12):2401-2410.
62. Zampieri C, Salarian A, Carlson-Kuhta P, Nutt JG, Horak FB. Assessing mobility at home in people with early Parkinson's disease using an instrumented Timed Up and Go test. Parkinsonism Relat Disord 2011;17(4):277-280.
63. Smith E, Walsh L, Doyle J, Greene B, Blake C. The reliability of the quantitative timed up and go test (QTUG) measured over five consecutive days under single and dual-task conditions in community dwelling older adults. Gait Posture 2016;43:239-244.

64. Mancini M, El-Gohary M, Pearson S, et al. Continuous monitoring of turning in Parkinson's disease: Rehabilitation potential. NeuroRehabilitation 2015;37(1):3-10.
65. El-Gohary M, Pearson S, McNames J, et al. Continuous monitoring of turning in patients with movement disability. Sensors (Basel, Switzerland) 2013;14(1):356-369.
66. Lord S, Godfrey A, Galna B, Mhiripiri D, Burn D, Rochester L. Ambulatory activity in incident Parkinson's: more than meets the eye? J Neurol 2013.
67. van Nimwegen M, Speelman AD, Smulders K, et al. Design and baseline characteristics of the ParkFit study, a randomized controlled trial evaluating the effectiveness of a multifaceted behavioral program to increase physical activity in Parkinson patients. BMC Neurol 2010;10:70.
68. Rochester L, Chastin SF, Lord S, Baker K, Burn DJ. Understanding the impact of deep brain stimulation on ambulatory activity in advanced Parkinson's disease. J Neurol 2012;259(6):1081-1086.
69. Chastin SFM, Granat MH. Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait & posture 2010;31(1):82-86.
70. Dontje ML, de Greef MH, Speelman AD, et al. Quantifying daily physical activity and determinants in sedentary patients with Parkinson's disease. Parkinsonism Relat Disord 2013;19(10):878-882.
71. Chastin SFM, Baker K, Jones D, Burn D, Granat MH, Rochester L. The pattern of habitual sedentary behavior is different in advanced Parkinson's disease. Movement Disorders 2010;25(13):2114-2120.
72. Cavanaugh JT, Ellis TD, Earhart GM, Ford MP, Foreman KB, Dibble LE. Capturing ambulatory activity decline in Parkinson's disease. Journal of neurologic physical therapy : JNPT 2012;36(2):51-57.
73. Cavanaugh JT, Ellis TD, Earhart GM, Ford MP, Foreman KB, Dibble LE. Toward Understanding Ambulatory Activity Decline in Parkinson Disease. Phys Ther 2015;95(8):1142-1150.
74. Lim I, van Wegen E, Jones D, et al. Does cueing training improve physical activity in patients with Parkinson's disease? Neurorehabil Neural Repair 2010;24(5):469-477.
75. Giuffrida JP, Riley DE, Maddux BN, Heldman DA. Clinically deployable Kinesia technology for automated tremor assessment. Mov Disord 2009;24(5):723-730.
76. Cancela J, Pastorino M, Tzallas AT, et al. Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors. Sensors (Basel, Switzerland) 2014;14(9):17235-17255.
77. Orendurff MS, Schoen JA, Bernatz GC, Segal AD, Klute GK. How humans walk: bout duration, steps per bout, and rest duration. Journal of rehabilitation research and development 2008;45(7):1077-1089.
78. Barry G, Galna B, Lord S, Rochester L, Godfrey A. Defining ambulatory bouts in free-living activity: Impact of brief stationary periods on bout metrics. Gait and Posture 2015;42(4):594-597.
79. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. The international journal of behavioral nutrition and physical activity 2015;12(1):159.
80. Taraldsen K, Chastin SFM, Riphagen II, Vereijken B, Helbostad JL. Physical activity monitoring by use of accelerometer-based body-worn sensors in older adults: A systematic literature review of current knowledge and applications. Maturitas 2012;71(1):13-19.
81. Storm FA, Heller BW, Mazza C. Step detection and activity recognition accuracy of seven physical activity monitors. PloS one 2015;10(3):e0118723.
82. Picerno P, Cereatti A, Cappozzo A. A spot check for assessing static orientation consistency of inertial and magnetic sensing units. Gait Posture 2011;33(3):373-378.

83. Godfrey A, Lara J, Munro CA, et al. Instrumented assessment of test battery for physical capability using an accelerometer: a feasibility study. Physiol Meas 2015;36(5):N71-83.
84. Cheng Z, Li P, Wang J, Guo S. Just-in-Time Code Offloading for Wearable Computing. Emerging Topics in Computing, IEEE Transactions on 2015;3(1):74-83.
85. Steins D, Sheret I, Dawes H, Esser P, Collett J. A smart device inertial-sensing method for gait analysis. J Biomech 2014;47(15):3780-3785.
86. Tsanas A, Little MA, McSharry PE, Ramig LO. Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. IEEE transactions on bio-medical engineering 2010;57(4):884-893.
87. Piro NE, Baumann L, Tengler M, Piro L, Blechschmidt-Trapp R. Telemonitoring of patients with Parkinson's disease using inertia sensors. Applied clinical informatics 2014;5(2):503-511.
88. Brouillette RM, Foil H, Fontenot S, et al. Feasibility, reliability, and validity of a smartphone based application for the assessment of cognitive function in the elderly. PloS one 2013;8(6):e65925.
89. Liddle J, Ireland D, McBride SJ, et al. Measuring the lifespace of people with Parkinson's disease using smartphones: proof of principle. JMIR mHealth and uHealth 2014;2(1):e13.
90. Vayena E, Tasioulas J. Adapting standards: ethical oversight of participant-led health research. PLoS medicine 2013;10(3):e1001402.
91. Kelly P, Marshall SJ, Badland H, et al. An ethical framework for automated, wearable cameras in health behavior research. American journal of preventive medicine 2013;44(3):314-319.
92. Heldman DA, Giuffrida JP, Chen R, et al. The modified bradykinesia rating scale for Parkinson's disease: reliability and comparison with kinematic measures. Mov Disord 2011;26(10):1859-1863.
93. Fisher JM, Hammerla NY, Rochester L, Andras P, Walker RW. Body-Worn Sensors in Parkinson's Disease: Evaluating Their Acceptability to Patients. Telemedicine journal and e-health : the official journal of the American Telemedicine Association 2016;22(1):63-69.
94. Prudon B, Duncan GW, Khoo TK, Yarnall AJ, Anderson KN. Primary sleep disorder prevalence in patients with newly diagnosed Parkinson's disease. Mov Disord 2014;29(2):259-262.
95. Bossenbroek L, Kosse N, Ten Hacken N, Gordijn M, Van der Hoeven J, De Greef M. Validation of the DynaPort MiniMod during sleep: a pilot study. Perceptual and motor skills 2010;111(3):936-946.
96. Godfrey A, Culhane KM, Lyons GM. Comparison of the performance of the activPAL Professional physical activity logger to a discrete accelerometer-based activity monitor. Med Eng Phys 2007;29(8):930-934.
97. Larkin L, Nordgren B, Purtill H, Brand C, Fraser A, Kennedy N. Criterion Validity of the ActivPAL Activity Monitor for Sedentary and Physical Activity Patterns in People Who Have Rheumatoid Arthritis. Phys Ther 2015.
98. Dijkstra B, Kamsma Y, Zijlstra W. Detection of gait and postures using a miniaturised triaxial accelerometer-based system: accuracy in community-dwelling older adults. Age Ageing 2010;39(2):259-262.
99. Kangas M, Vikman I, Wiklander J, Lindgren P, Nyberg L, Jamsa T. Sensitivity and specificity of fall detection in people aged 40 years and over. Gait Posture 2009;29(4):571-574.
100. Pansera M, Estrada JJ, Pastor L, Cancela J, Greenlaw R, Arredondo MT. Multi-parametric system for the continuous assessment and monitoring of motor status in

Parkinson's disease: an entropy-based gait comparison. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2009;2009:1242-1245.
101. Salarian A, Russmann H, Vingerhoets FJ, Burkhard PR, Aminian K. Ambulatory monitoring of physical activities in patients with Parkinson's disease. IEEE transactions on bio-medical engineering 2007;54(12):2296-2299.
102. Bouten CV, Verboeket-van de Venne WP, Westerterp KR, Verduin M, Janssen JD. Daily physical activity assessment: comparison between movement registration and doubly labeled water. Journal of applied physiology (Bethesda, Md : 1985) 1996;81(2):1019-1026.
103. Benka Wallen M, Franzen E, Nero H, Hagstromer M. Levels and Patterns of Physical Activity and Sedentary Behavior in Elderly People With Mild to Moderate Parkinson Disease. Phys Ther 2015;95(8):1135-1141.
104. Peterson NE, Sirard JR, Kulbok PA, DeBoer MD, Erickson JM. Validation of Accelerometer Thresholds and Inclinometry for Measurement of Sedentary Behavior in Young Adult University Students. Research in nursing & health 2015;38(6):492-499.
105. White DK, Wagenaar RC, Ellis T. Monitoring activity in individuals with Parkinson disease: a validity study. Journal of neurologic physical therapy : JNPT 2006;30(1):12-21.
106. Schmidt AL, Pennypacker ML, Thrush AH, Leiper CI, Craik RL. Validity of the StepWatch Step Activity Monitor: preliminary findings for use in persons with Parkinson disease and multiple sclerosis. Journal of geriatric physical therapy (2001) 2011;34(1):41-45.
107. Galna B, Lord S, Burn DJ, Rochester L. Progression of gait dysfunction in incident Parkinson's disease: impact of medication and phenotype. Mov Disord 2015;30(3):359-367.